

Rust in Peace: Psychological Methods for Detecting Bots

Christopher T. Rodriguez & Daniel M. Oppenheimer

Carnegie Mellon University, Department of Social & Decision Sciences

Email: crodrig3@andrew.cmu.edu

Abstract

- Bots are getting better at completing standard CAPTCHAs which has dire consequences for quality of data collected via online crowdsourcing platforms.
- In our sample of 906 MTurkers, we identify 171 bots, while ReCAPTCHA V3 identifies 16 bots.
- We propose a variety of psychological methods aimed to help screen out bots.

Motivation

- The presence of bots in online crowdsourcing platforms lowers the data quality collected on those platforms.¹
- Advances in AI have hindered the performance of CAPTCHAs as a sole defense against bots.²
- Open-ended free response questions have been shown to be effective bot identifiers.³
- **It is our goal to develop better automated bot screeners.**

Methods

- We recruited 906 MTurkers to take a survey that consisted of demographics, a CAPTCHA, identification questions, and bot-screening questions.
- Identification questions were free response questions that required causal reasoning and image processing (e.g., Figure 1). These questions were manually coded to identify participants as human or bot.
- Bot-screening questions aimed to be difficult for bots and easy for humans. We created 7 categories of these questions: Sensory, Learning, Psychometric, Theory of Mind, Identify-Sort-Add, Image Processing & Causal Reasoning, and Personal References.

Figure 1: What would happen to the feet pictured below on a warm, sunny day?



The above question is an example of one of the four identification questions that participants had to complete in our survey. Responses to these questions were later categorized as bot or human. Bots struggle with these questions (e.g., giving "SNOW" as an answer), whereas humans identify that the feet will melt.

Analysis

- Three independent RAs coded ID response as bot (0), human (1), or unsure (.5). These ratings were summed to form a "bot score."
- By categorizing unanimous bots (bot score = 0) and unanimous humans (bot score = 12), we could examine how well our bot screeners were able to discern bots from humans.
 - It is possible that some participants labelled "unanimous bots" could be humans that are inattentive, lazy, and/or, non-English speaking participants, we believe this distinction matters little. Ultimately, these participants would also lower data quality, so knowing what attention checks will screen them out is valuable.
- We identified **171 "unanimous bots"** (all three RAs independently designating each of the ID responses as bot, or a total bot score of 0).
 - ReCAPTCHA V3 identified **merely 16 potential bots**.

Table 1: Unanimous Bot and Unanimous Human Performance on Attention Checks

Question Type	Unanimous Bots	Unanimous Humans
Sensory A*	48.8%	99.7%
Sensory B*	36.3%	93.2%
Learning	25.2%	83.8%
Psychometric A	45.0%	99.2%
Psychometric B	65.5%	97.9%
Psychometric C	31.6%	69.5%
Theory of Mind A*	32.5%	93.1%
Theory of Mind B	9.4%	83.1%
Identify-Sort-Add	10.8%	87.1%
Image Processing & Causal Reasoning A*	51.2%	98.4%
Image Processing & Causal Reasoning B*	54.6%	99.5%
Personal References*	38.3%	98.4%

This chart displays the correct answer rate for bots and humans for each question type. * Indicates the questions were multiple choice with four answer choices.

Results & Discussion

- Humans outperformed bots on each of our bot screeners, typically by large margins. All our bot checks outperformed ReCAPTCHA V3 for identifying bots.
- Bots performed above chance for multiple choice questions, but these questions can be made more difficult by expanding choice sets (personal references could be improved this way).
 - These two-part questions require internal consistency from participants. The first question asks for an opinion. The second question contains a correct answer dependent on the first.
- Notably, bots seemed to struggle during tasks that required perspective taking.
- As we've seen with CAPTCHAs, it will only be a matter of time until bots' ability advances. It is our aim to provide bot checks you can use in online surveys as well as inspire creation of future checks in a never-ending arms race against bots.

References

¹Godinho, A., Schell, C., & Cunningham, J. A. (2020). Out damn bot, out: Recruiting real people into substance use studies on the internet. *Substance Abuse*, 41(1), 3–5.

²Sivakorn, S., Polakis, J., & Keromytis, A. D. (2016). I'm not a human: Breaking the Google reCAPTCHA. *Black Hat*, 14.

³Storozuk, A., Ashley, M., Delage, V., & Maloney, E. A. (2020). Got bots? Practical recommendations to protect online survey data from bot attacks. *The Quantitative Methods for Psychology*, 16(5), 472–481